White Paper Report

Report ID: 110231

Application Number: HT-50080-13

Project Director: Clifford Anderson (clifford.anderson@vanderbilt.edu)

Institution: Vanderbilt University

Reporting Period: 10/1/2013-12/31/2014

Report Due: 3/31/2015

Date Submitted: 3/30/2015

White Paper

Grant Number: HT-50080-13

**XQuery Summer Institute:
Advancing XML-Based Scholarship
from Representation to Discovery**

Project Director: Clifford Anderson

Vanderbilt University

March 30, 2015

Executive Summary

The XQuery Summer Institute was an National Endowment for the Humanities (NEH)-funded Institute for Advanced Topics in the Digital Humanities. The Institute took place on the campus of Vanderbilt University from June 9 to 20, 2014. The goal of the Institute was to assist digital humanists from diverse locations in the academy to work computationally with documents marked up in XML. Our aspiration was to help researchers to make the transition from encoding documents in markup languages to exploring those documents computationally. The Institute proved very successful: participants reported a 97.17% satisfaction rate with its overall performance. All participants reported six months after the conclusion of the Institute that they are using or planning to use XQuery for digital projects.

Project Activities

The period of the grant was October 2013 to December 2014. The primary activity of the XQuery Summer Institute was hosting a two-week session on XQuery at Vanderbilt University from June 9 to 20, 2014. Six project faculty instructed a cohort of twelve digital humanists and librarians in the fundamentals of XQuery, a domain-specific programming language for XML, and eXist, an open source XML database. The Institute taught participants how to build web-based digital applications to browse and search texts encoded in TEI. The months leading up to the Institute were spent primarily on advertising, selecting participants, logistical arrangements, and developing curriculum. In the months following the Institute, we kept in touch with participants through our listserv, bimonthly Google Hangouts, and by sending out XQuery exercise problems.

The project faculty were Clifford Anderson, Director for Scholarly Communications at Vanderbilt University, and Project Leader of the Institute; Kevin S. Clarke, Digital Library Programmer at the University of California Los Angeles; David Michelson, Assistant Professor of The History of Christianity and Classical Studies at Vanderbilt University; Dale Poulter, Coordinator of Search and Core Services at Vanderbilt University; Jonathan Robie, the Lead Editor of the XQuery and XPath specifications at the World Wide Web Consortium (W3C); and Winona Salesky, an independent digital library consultant.

Audiences

The primary audience for the XQuery Summer Institute was digital humanists, including "alt-ac" professionals, faculty members, graduate students, and librarians, who had some experience working with XML-encoded metadata but who had not yet learned to work with that metadata computationally. The proposal for the Institute was based on our anecdotal perception that many digital humanists learn how to encode texts in computer-readable markup languages without learning how to query and analyze them computationally. In our grant proposal, we proposed to select twelve participants in roughly this position and help them to take the next step by learning to program against their data using XQuery.

Publicity

The publicity for the XQuery Summer Institute was straightforward and simple. The three primary media for advertising the Institute were a website, various listservs, and Twitter. The goal of our publicity was to ensure as diverse a group of applicants as possible.

The website was created using GitHub Pages. GitHub offers free web hosting for open source organizations on its platform. The only cost associated with hosting the website was the registration of its domain name, XQueryInstitute.org. The website went live in October 2013.

After the launch of the website, the project leader and his Vanderbilt colleagues emailed a standard announcement to the major digital humanities websites, including Art Libraries Society of North America Digital Humanities Special Interest Group, Digital Humanities Now, HASTAC, NYC Digital Humanities and the TEI Listserv, among other places.

A key channel for advertising the Institute was Twitter. The various project faculty tweeted regularly about the upcoming Summer Institute to their followers. We also benefitted from tweets and retweets by the staff members at the Office of Digital Humanities at the NEH.

Another significant source of publicity in the lead up to the Institute was an interview between Jason Rhody and Clifford Anderson published on February 10, 2014 on the NEH website: ODH Project Q&A: the XQuery Summer Institute. The opportunity to explain the technology in context helped broaden our pool of applicants for the Institute.

In the week prior to the beginning of the Institute, Vanderbilt University published an article titled Vanderbilt to host NEH-supported institute on digital humanities. (See Appendix Two for the announcement.)

Participants were encouraged to tweet actively during the Institute. In fact, one of our first items of business was deciding on our hashtag: #xqy14. The 'twitterstream' from the event provided a window for other digital humanists into the Institute's activities.

Participants

We received 37 applications for 12 spots at the Institute. A panel of three (Clifford Anderson, Dave Michelson, and Dale Poulter) evaluated the applications. As stated in the grant proposal, we favored those who: 1. had prior experience with XML-based markup languages; 2. applied with team members from the same institution; 3. came from diverse backgrounds in the academy.

We expected candidates to have had some prior experience working with some form of XML metadata, though we did not specify experience with any particular kind of metadata. Our hope was to identify scholars who were already encoding metadata in some form and assist them to make the transition to exploring their data computationally. Though we reviewed the basics of XML (and related technologies like XML schemas) during the opening day, the Institute assumed a baseline knowledge of markup languages. In our pre-participation survey, most indicated that they had an intermediate level of experience with

XML, but had little-to-no knowledge of XQuery and eXist. When asked about their experience 91.7% of our participants also reported that they were currently working on XML-based projects.

The Institute did not assume any experience with programming languages. However, our pre-participation survey revealed that the majority of our participants (83.3%) already had some programming experience. Of those arriving to the Institute with a programming background, most had prior experience with JavaScript (60%) and XSLT (50%). Other popular programming language choices included PHP and Python. The difference in programming background made for an uneven playing field among participants. While we avoided making analogies to other programming languages when teaching XQuery (along the lines of "X in XQuery is equivalent to Y in SQL," etc.), participants who already understood programming concepts like functions, data types, and error handling arrived at an advantage to those who had never programmed before. In particular, participants with a strong grasp of XSLT had a good head start on others because XSLT shares XPath and the XML Data Model (XDM) in common with XQuery. Participants who had programmed in other languages like PHP and Python needed to come to terms with differences between those imperative and object-oriented languages and functional programming languages like XQuery.

Having anticipated such differences in participants' backgrounds, we made use of pair programming throughout the Institute. We encouraged participants to work together in pairs, alternating between keyboardist and observer roles. We adopted this practice from the Extreme Programming community, which uses pair programming to mentor junior developers and also to improve code quality. Our hope was that team members who applied from the same institution might continue to collaborate in this manner after returning from the Institute.

Finally, we aspired to gather a diverse group of participants from different areas of the academy. In response to the question "How do you identify yourself professionally?" on our pre-participation survey, candidates selected Librarian (50%), Faculty (33.3%), Digital Humanist (25%); Archivist (16.7%) and "Other" (8.3%). (The survey allowed more than one response to that question.) The diversity of our participants' perspectives on XQuery in the digital humanities made for a stimulating and exciting learning environment.

All participants were notified of decisions by March 18, 2014.

Institute Performance

The Institute took place over two weeks. The first week was dedicated to learning the fundamentals of the XQuery language while the second week was devoted to using XQuery and the eXist XML database to build web-based applications with XQuery.

By and large, we followed the curriculum we initially proposed in our application. However, we took the opportunity in several places to capitalize on outside experts. We also took the advice of the NEH staff to provide opportunities for all participants to speak about their own digital humanities projects. (See Appendix One for the final curriculum.)

Our guest speakers provided expert commentary on key topics. For example, we invited Rebecca Niles, Digital Editions Editor, and Michael Poston, Database Applications Associate, at the Folger Shakespeare Library to discuss their open source editions of Shakespeare via a Skype call with the Institute participants on Tuesday, June 10. We benefited greatly from the chance to learn more about the genesis of those digital texts and to ask questions about various encoding decisions. We returned repeatedly to topics raised during our conversation with Niles and Poston as we explored the Folger edition of *Julius Caesar* during the Institute. We also invited Jay Clayton, William R. Kenan, Jr. Professor, Department of English at Vanderbilt University, to speak to us about the different modalities of the digital humanities. Having his broad perspective on the discipline and the place of digital text encoding in it helped us to even the playing field for participants who were relatively new to the digital humanities.

We also allowed the project faculty to tailor their instructional sessions as they saw fit. In most cases, changes were minor, primarily reflecting the need to pace the instruction appropriately to ensure that everyone could follow along. For example, we added an extra session on Wednesday, June 18 so Dale Poulter could lead us through the essentials of setting up accounts on Amazon Web Services (AWS) before our final deployments that Friday.

At the suggestion of the NEH staff, we provided opportunities for all Institute participants to present about their own digital projects. These presentations added richness and depth to our program. We learned a great deal from each other's projects and activities. The participants' talks also allowed our project faculty to make suggestions about how best to apply XQuery (and related technologies) in the participants' particular contexts.

Dave Michelson and I taught the first two days of the Institute. These days consisted largely of review of XML and TEI fundamentals. Our goal was to make sure that everyone started learning XQuery with more or less the same understanding of the XML data that we would be exploring. However, we found that the material was new for some participants. In retrospect, we might have assigned some of this material prior to the beginning of the Institute as practice. However, covering these topics during the Institute itself gave us a chance to gauge the experience of our participants, to answer questions, and to learn from each other. We also discovered, as may be expected, that we all encoded TEI documents very differently depending on our projects.

We progressed a little more slowly than we anticipated in the three days that Jonathan Robie instructed us in XQuery fundamentals, primarily because we delved into many of the concepts at a deeper level. We took full advantage of Robie's vast knowledge of XQuery. As a co-editor of the XQuery Recommendation, Robie provided us with unique insight into the development of the language standard. During the second week, I covered some of the advanced topics that we passed over during the first week at impromptu lunch gatherings. An advantage of these impromptu sessions is that they allowed participants to decide whether or not to participate depending on their understanding of the more basic sessions.

In the second week, we introduced the eXist database. We aspired to provide participants with hands-on experience building a web application with XQuery, eXist, and XML datasets.

Winona Salesky arrived on Monday to introduce eXist, an open source XML database. She illustrated her concepts by reference to the sample Shakespeare application she had built prior to the Institute. Kevin Clarke arrived on Wednesday to teach key concepts in software engineering such as designing REST web services with XQuery and implementing Model-View-Controller architectures in eXist. Dale Poulter assisted participants with deploying the sample Shakespeare application to Amazon Web Services. Some glitches marred the deployment, though at least one participant managed to get the application up and running. In retrospect, we did not allot sufficient time to moving the application to the Amazon cloud, which required understanding IT concepts like configuring servers, forwarding ports, and setting up adequate firewalls. Attempting to cover these concepts, even in a nutshell, was probably too ambitious at the conclusion of two weeks of intense instruction.

The differences between digital humanists and librarians became somewhat more apparent during the second week. In general, librarians approach digital projects as a means of providing broad access to intellectual resources whereas digital humanists tend toward research-driven agendas. Most librarians (and those working in academic or educational technology) aspire to build interactive digital systems. In general, they have a stronger background in information technology. By contrast, many digital humanists have stronger research interests in their datasets. Their primary goal is learning techniques for exploratory data analysis to support (or disprove) research hypotheses. Building interactive systems is, to some extent, incidental to their reasons for learning XQuery.

While by no means absolute, this divide shaped the different assumptions and goals of participants. A key challenge during the second week was to satisfy, to the extent possible, both groups while covering fairly technical topics in software engineering.

Finally, we added two receptions to the schedule. The University Library hosted a reception on June 9. Connie Dowell, Dean of Libraries, addressed participants at a wine and cheese reception, to which all campus librarians and library staff members were invited. On June 18, the Robert Penn Warren Center for the Humanities hosted a reception at its historic campus home. Mona Frederick, Executive Director of the Center, invited Vanderbilt faculty members and graduate students working in the field of digital humanities to the reception. These events provided informal opportunities for participants and members of the local digital humanities community to meet and exchange ideas.

Evaluation

Our primary method of evaluation was three surveys. Participants were surveyed on the opening day of the Institute, immediately after the Institute, and six months after the conclusion of the Institute. The surveys indicate high levels of satisfaction with the Institute.

In the immediate post-participation survey, the mean score of participants when asked about their overall ranking of the Institute was 97.17% and the median score was 100%. All of the participants answered "Yes" when asked whether they would recommend the XQuery Summer Institute to friends and colleagues.

The mean participant score when asked about their knowledge of XML in the pre-participation survey was 50.17%, which we described as "intermediate" on our survey labels. (We substituted 50% for all NAs in our dataset on the assumption that participants meant to indicate intermediate-level knowledge if they did not move the slider off its initial setting of 50% in both surveys.) The mean participant score when responding to the same question in the post-participant survey was 58.7%, a modest increase of approximately 8.5%. This accords with our assumptions that most participants would have an intermediate level of knowledge of XML at the beginning of the Institute.

The participants' perceived understanding of XQuery grew more substantially over the course of the two weeks. At the beginning of the Institute, participants' mean score when asked about their knowledge of XQuery was 4.5% or what we labelled as "Beginner" in our survey labels. By the conclusion of the Institute, participants' mean rating of their knowledge was 38.17%, a significant increase of 33.67%. Given the challenge of learning any programming language, it's not surprising that nearly all participants ranked themselves somewhere between "beginner" and "intermediate" after the conclusion of the Institute. A two week course in any programming language, no matter how intensive, is sufficient only to get participants started on the road to developing genuine expertise.

In our proposal, we stated that we would judge our success primarily on the basis of a survey six months after the Institute. All participants completed the six-month survey, though in some case they did so more than six months later. The survey results showed continuing enthusiasm for the potential of XQuery. All respondents answered "Yes" to the question "Do you plan to use or continue to use XQuery for scholarly projects?" 50% responded that they are currently working on XQuery-related projects.

In response to the question, "Did you use XQuery during the past six months?" 75% responded "Yes" and 25% responded "No." As for hours of practice using XQuery, participants reported on the post-participation survey that they anticipated working on XQuery-related projects for 2 to 4 hours (66.7%) or 5 to 7 (33.3%) hours per week during the next six months. Of those actively working using XQuery after six months, they reported working 0 to 1 (44.4%) or 2 to 4 (55.6%) hours per week. While participants' expectations were a bit chastened by the reality of competing demands on their time, these results indicate ongoing appreciation for the potential of XQuery in the digital humanities and librarianship.

Continuation of the Project

While the XQuery Summer Institute was a one-time event made possible by NEH funding, we hope to continue building the community of XML and XQuery enthusiasts in the academy. Right now, the community is fairly small. If, as we foresee, XQuery will become a mainstay of digital humanities programming, we will need to develop better and more targeted teaching tools. Our experience teaching XQuery to participants in the Institute and other contexts helped us to identify certain lacuna in the teaching material and approach.

First, and most crucially, we need teaching materials aimed at digital humanists (and librarians). Initially, we had assigned three texts for the Institute: Joe Fawcett, et. al., eds.

Beginning XML, Fifth Edition (Indianapolis: Wrox, 2012), Priscilla Walmsley's XQuery (Sebastopol, CA: O'Reilly, 2007), and Ron Hitchens' Getting Started with XQuery: Querying XML Like You Mean It (Dallas: Pragmatic Press, 2008). The Wrox book on Beginning XML was meant to cover the gaps in background knowledge and to serve as a collective point of reference. Walmsley's text is the definitive introduction and reference work for XQuery. Hitchens provided an accessible and step-by-step introduction to XQuery and two of the most popular XML databases: eXist and MarkLogic. Unfortunately, our participants alerted us in the lead up to the Institute that Hitchens' book was no longer available from the publisher. We assume that Pragmatic Press removed the book from its list because its coverage of eXist and MarkLogic was no longer current. However, the unavailability of that book meant that we no longer had an accessible introduction to assign for the Institute. While Walmsley covers XQuery in exemplary fashion, she writes primarily for professional programmers and does not cover the database implementations in detail.

Second, we need to adapt our teaching of programming to the common patterns of digital humanities research. Here we may learn from the R community. R is an open source statistical programming language used for data science. Many users of R are not programmers in any traditional sense. Rather, they are domain experts with a need for statistical analysis. The R community has developed tools to enable researchers to work effectively with their data with only a minimum of programming skills. In particular, many researchers use the R Read-Evaluate-Print-Loop (REPL) to explore their datasets by issuing one command at a time. Their understanding of R as a programming language grows slowly over time as their requirements grow increasingly complex. Our teaching of XQuery, especially during the first week of the Institute, suggests that much the same pattern is applicable for scholars working on textual editions in the digital humanities. As Jonathan Robie underscored during his instructional sessions, encoding and querying should not be regarded as distinct phases in digital projects. Rather, we should adopt an iterative approach, encoding some data and then testing the suitability of that encoding for various purposes by exploring and reshaping it computationally with XQuery. Analogously, we should not leave encoding to domain experts and writing queries to the "programmers." Following the model set by the R community, domain experts should learn to write queries *as* they encode and explore TEI datasets. This insight has the potential to change the teaching of TEI and related XML standards in the digital humanities.

Finally, we need to take a more expansive perspective on what constitutes computational literacy. Currently, we teach primarily how to *write* programming code. We spend very little time teaching digital humanists how to read and speak about programming languages. Yet digital humanists frequently find themselves in the position of reading and speaking about code written by collaborators. When we propose to teach humanists to program, we may want to adopt a broader vision of computational literacy. Does literacy always mean learning to write code? It may be that speaking and reading code are more crucial skills for humanists when collaborating with programmers. We may not wish, therefore, only to judge the success of events like the XQuery Institute on the lines of code written by participants. If participants have become more confident in speaking and reading code, we will have contributed significantly to the partnerships between technologists and humanists that propel the digital humanities.

Long-Term Impact

The community of digital humanists along with funding agencies like the Office of Digital Humanities at the NEH have made significant investments in XML-based research and scholarship. Digital humanists have a vested interest in the future of XML since it's uniquely adapted for the description of narrative documents. While alternative data standards like JavaScript Object Notation (JSON) have gained popularity among the current generation of web programmers, XML and its related technologies remain the most appropriate data standards for narrative description. In particular, the Text Encoding Initiative (TEI) standard is unsurpassed as a data standard for textual encoding in the humanities. While technically possible, encoding a critical edition in JSON would prove an exercise in frustration. XML will likely remain a standard of humanities computing for many years to come.

We hope that there will be growing interchanges between the XML Activity at the W3C, industry (especially digital publishers), and the digital humanities community. The flourishing of XML-based technologies redounds to the benefit of digital humanists and publishers. This is particularly true with respect to open source XML technologies like BaseX and eXist, among others. We believe that, in its small way, the XQuery Summer Institute brought together these communities. We look forward to continuing these kinds of collaborations among different groups of XML users and stakeholders.

Grant Products

The GitHub repository for the XQuery Institute contains the syllabus and many instructional materials from the Institute. These materials are licensed under a Creative Commons (CC-BY 3.0) license, making them widely reusable. The repository also contains the three follow up exercises sent in the months following the conclusion of the Institute.

We invited participants in the Institute to bring XML datasets with them. However, we believed it necessary to adopt a common dataset and metadata standard for pedagogical purposes during the Institute. We adopted the Folger Digital Texts of Shakespeare. These documents are encoded in TEI and released under a Creative Commons Attribution-NonCommercial 3.0 Unported license.

During the second week of the Institute, participants took part in building a web application in XQuery for the Folger Digital Texts. Winona Salesky was the primary author of the web application. Kevin S. Clarke wrote the build scripts for the application. The source code for this demonstration application is available on GitHub and has been released as open source under the MIT License. After the completion of the Institute, we alerted the staff members at the Folger Shakespeare Library about the availability of the demo application. We hope that the Folger Library and other digital humanists will build on this application and reuse it in other teaching contexts.

Finally, I presented a paper dealing in part with the XQuery Institute at *Balisage: The Markup Conference* in Washington, DC on August 14, 2014. The presentation summarized the goals of the Institute and provided an opportunity for feedback and suggestions from

leading experts in the XML world. The paper, titled On Teaching XQuery to Digital Humanists, is published in the *Proceedings of Balisage: The Markup Conference 2014* under a Creative Commons Attribution 3.0 license.

Acknowledgements

Staging the XQuery Summer Institute required assistance from many hands. In addition to our project faculty and participants, I owe a large debt of thanks to many others. First, I'd like to thank Jason Rhody, Senior Program Officer at the Office of Digital Humanities at the NEH. I benefited greatly from his prompt and helpful advice at all stages of the project. I would also like to thank Jennifer Serventi, Senior Program Officer at the Office of Digital Humanities, for conducting a site visit during the final days of the Institute. The participants, project faculty, and I deeply appreciated the NEH's commitment to make the Institute successful.

I would like to thank Joseph D. Combs, Jr., Interim Dean of Libraries as well as Connie Vinita Dowell, former Dean of Libraries, for their support of the Institute. I am grateful to Jean Klockenkemper, Executive Director of Finance and Administration at the Vanderbilt University Library, for guiding me through the budgeting process. Nancy Dwyer, Programs and Special Projects Coordinator at the Vanderbilt University Library, helped with all aspects of organization. She arrived every morning and afternoon of the Institute with refreshments. Crucially, she also made sure that leaders were paid and that participants received their reimbursements. Susan Grider, Administrative Coordinator in the Vanderbilt Law Library, graciously put packets together for participants. Ed Warga, Institutional Repository Coordinator at the Vanderbilt University Library, assisted in many ways in the lead up to the conference, including by proofreading our curricular materials.

I am grateful to Mona Frederick, Executive Director of the Robert Penn Warren Center at Vanderbilt, for her hospitality to our participants and faculty during their visit to Vanderbilt. My thanks also to the staff at the Vanderbilt University Office of Conferences for organizing the housing and meals during the Institute: E. Michelle Wu, Conference and Event Manager, H. G. Stovall, former Assistant Director of Conferences, and Jessica Posey, Program Coordinator. Finally, I would like to express my appreciation to Clint Brown, Director, and Lisa Au, Sponsored Program Specialist, in the Office of Sponsored Programs at Vanderbilt University for assistance with the grant writing and submission process.

Survey data from the XQuery Summer Institute participants was collected using REDCap, supported by UL1 TR000445 from NCATS/NIH.

Appendix One
XQuery Summer Institute: Detailed Schedule and Curriculum

Sunday, June 8

Check-in (4 p.m. to 9 p.m.)

Lewis House, Highland Quadrangle on the Vanderbilt University campus

N.B. The Lewis House reception desk will be open until 12 a.m. If you arrive after midnight, please call (615) 343-8699 for assistance.

Monday, June 9

Breakfast (8:00 a.m. in Rand Hall)

Morning Session I (9:00 a.m. to 10:15 a.m.)

Instructor: Clifford B. Anderson

- Introduction to the Institute and participants
- Discussion of the Institute's goals and anticipated outcomes
- Assignment of initial pair programming partnerships

Morning Session II (10:30 a.m. to 11:45 a.m.)

Instructor: David A. Michelson

- Review of XML essentials, including namespaces, character entities, and CDATA sections

- Session Notes

Readings:

- Joe Fawcett, Liam Quin, and Danny Ayers, *Beginning XML*, Chapter 1: What is XML?, Chapter 2: Well-Formed XML, and Chapter 3: XML Namespaces

- David J. Birnbaum, "What is XML and why should humanists care? An even gentler introduction to XML."

Lunch (11:45 a.m. in Rand Hall)

Afternoon Session I (1:00 p.m. to 2:15 p.m.)

Instructor: Clifford B. Anderson

- Review of Document Type Definitions (DTDs) and XML Schemas
- A brief look at Relax NG and Schematron

Readings: Joe Fawcett, Liam Quin, and Danny Ayers, *Beginning XML*, Chapter 4: Document Type Definitions; Chapter 5: XML Schemas; Chapter 6: Relax NG and Schematron

Afternoon Session II (2:30 p.m. to 3:45 p.m.)

Workshop Format

- Installing eXist, oXygen, GitHub and loading of Folger Shakespeare Texts.

Reception (4:00 p.m. to 5:00 p.m.)

The University Library will host a reception for our Institute in the 2nd Floor Gallery.

Dinner (5:45 p.m. at The Commons)

Tuesday, June 10

Breakfast (8:00 a.m. in Rand Hall)

Morning Session I (9:00 a.m. to 10:15 a.m.)

Instructor: David A. Michelson

- The challenges of scholarly editing in XML
- Session notes

Readings:

- *Browse*: Lou Burnard, Katherine O'Brien O'Keeffe, and John Unsworth, eds. *Electronic Textual Editing*

Morning Session II (10:30 a.m. to 11:45 a.m.)

Instructor: David A. Michelson

- Introduction to the TEI P5 specification
- Session notes

Readings:

- P5: Guidelines for Electronic Text Encoding and Interchange, Chapter 3: Elements Available in All TEI Documents

- *Browse*: Lou Burnard, *What is the Text Encoding Initiative?*

Lunch (11:45 a.m. in Rand Hall)

Afternoon Session I (1:00 p.m. to 2:15 p.m.)

Skype Conversation with Michael Poston & Rebecca Niles

Folger Shakespeare Library, Washington, D.C.

- Exploration of the Folger Library's digital edition (TEI P5) of Julius Caesar

Afternoon Session II (2:30 p.m. to 3:45 p.m.)

Instructor: Dr. Jay Clayton, William R. Kenan, Jr. Professor, Department of English, Vanderbilt University

- Transmediation and the Production of Digital Editions

Open session (4:00 p.m. to 5:00 p.m.)
- Hopes, aspirations, and collective goals for XQuery

Dinner (5:45 p.m. at The Commons)

Wednesday, June 11

Breakfast (8:00 a.m. in Rand Hall)

Morning Session I (9:00 a.m. to 10:15 a.m.)

Instructor: Jonathan Robie

- Introduction to XQuery

Readings: Priscilla Walmsley, *XQuery*, Chapter 1: Introduction to XQuery and Chapter 2: XQuery Foundations

Morning Session II (10:30 a.m. to 11:45 a.m.)

Instructor: Jonathan Robie

- Path expressions

Reading: Joe Fawcett, Liam Quin, and Danny Ayers, *Beginning XML*, Chapter 7: Extracting Data from XML; Priscilla Walmsley, *XQuery*, Chapter 4: Navigating Input Documents Using Paths

Lunch (11:45 a.m. in Rand Hall)

Afternoon Session I (1:00 p.m. to 2:15 p.m.)

Instructor: Jonathan Robie

- FLWOR expressions

Reading: Priscilla Walmsley, *XQuery*, Chapter 6: Selecting and Joining Using FLWORs

Afternoon Session II (2:30 p.m. to 3:45 p.m.)

Instructor: Jonathan Robie

- Constructors

Reading: Priscilla Walmsley, *XQuery*, Chapter 5: Adding Elements and Attributes to Results

Open session (4:00 p.m. to 5:00 p.m.)

- Exploring participants' datasets

Dinner (5:45 p.m. at The Commons)

Thursday, June 12

Breakfast (8:00 a.m. in Rand Hall)

Morning Session I (9:00 a.m. to 10:15 a.m.)

Instructor: Jonathan Robie

- Functions and modules

Morning Session II (10:30 a.m. to 11:45 a.m.)

Instructor: Jonathan Robie

- Functions and operators
    - partial function application
    - Inline functions
    - A brief look at second order functions

Lunch (11:45 a.m. in Rand Hall)

Afternoon Session I (1:00 p.m. to 2:15 p.m.)

Instructor: Jonathan Robie

- Grouping

Afternoon Session II (2:30 p.m. to 3:45 p.m.)

Institute Faculty

- Exploring participants' datasets

Open session (4:00 p.m. to 5:00 p.m.)

- Commentary, discussion, questions

Dinner (5:45 p.m. at The Commons)

Friday, June 13

Breakfast (8:00 a.m. in Rand Hall)

Morning Session I (9:00 a.m. to 10:15 a.m.)

Instructor: Jonathan Robie

- Windowing

Morning Session II (10:30 a.m. to 11:45 a.m.)

Instructor: Jonathan Robie

- XQuery spare change
    - if..then..else
    - effective boolean values
    - try..catch
    - switch expressions
    - typeswitch
    - quantifiers
    - mapping operator
    - string concatenation operators

Reading: Priscilla Walmsley, *XQuery*, Chapter 3: XQuery Building Blocks

Lunch (11:45 a.m. in Rand Hall)

Afternoon Session I (1:00 p.m. to 2:15 p.m.)

Instructor: Jonathan Robie

- Recursion and avoiding recursion

Afternoon Session II (2:30 p.m. to 3:45 p.m.)

Institute Faculty

- Using typeswitch transformations to render Shakespeare as XHTML

Open session (4:00 p.m. to 5:00 p.m.)
- Commentary, discussion, questions

Monday, June 16

Breakfast (8:00 a.m. in Rand Hall)

Morning Session I (9:00 a.m. to 10:15 a.m.)

Instructor: Winona Salesky

- Exploring eXist - An Open Source XML Database

Reading: Joe Fawcett, Liam Quin, and Danny Ayers, *Beginning XML*, Chapter 10: XML and Databases

Morning Session II (10:30 a.m. to 11:45 a.m.)

Instructor: Winona Salesky

- Combining eXist's extension functions with FLWOR expressions

Lunch (11:45 a.m. in Rand Hall)

Afternoon Session I (1:00 p.m. to 2:15 p.m.)

Instructor: Winona Salesky

- Exploring eXist's indexing options

Afternoon Session II (2:30 p.m. to 3:45 p.m.)

Institute Faculty

- Conducting full-text searches across Shakespeare's corpus using eXist's extension functions

Open session (4:00 p.m. to 5:00 p.m.)
- Participant presentation

Dinner (5:45 p.m. in Rand Hall)

Tuesday, June 17

Breakfast (8:00 a.m. in Rand Hall)

Morning Session I (9:00 a.m. to 10:15 a.m.)

Instructor: Winona Salesky

- Writing user defined functions
- Function signatures
- Checking types

Reading: Priscilla Walmsley, *XQuery*, Chapter 8: Functions

Morning Session II (10:30 a.m. to 11:45 a.m.)

Instructor: Clifford B. Anderson

- Handling errors in XQuery
- Strategies for diagnosing and preventing errors
    - XQLint

- XQDoc

Lunch (11:45 a.m. in Rand Hall)

Afternoon Session I (1:00 p.m. to 2:15 p.m.)

Instructor: Winona Salesky

- Developing modules of user defined functions in eXist

Reading: Priscilla Walmsley, *XQuery*, Chapter 12: Queries, Prologs, and Modules

Afternoon Session II (2:30 p.m. to 3:45 p.m.)

Institute Faculty

- Creating facets for searches

Open session (4:00 p.m. to 5:00 p.m.)
- Participant presentation

Dinner (5:45 p.m. in Rand Hall)

Wednesday, June 18

Breakfast (8:00 a.m. in Rand Hall)

Morning Session I (9:00 a.m. to 10:15 a.m.)

Instructor: Kevin Clarke

- Reviewing HTTP
- A brief introduction to REST

Reading: Cornelia Davis, "What if the web were not RESTful?" In Proceedings of the Third International Workshop on RESTful Design (WS-REST '12), Rosa Alarcon, Cesare Pautasso, and Erik Wilde (Eds.). New York: ACM, 3-10

Morning Session II (10:30 a.m. to 11:45 a.m.)

Instructor: Winona Salesky

- Introduction to eXist's httpclient extension functions

Lunch (11:45 a.m. in Rand Hall)

Afternoon Session I (1:00 p.m. to 2:15 p.m.)

Institute Faculty

- Connecting with online data sources using HTTP

Reading: Charles Severance, "Discovering JavaScript Object Notation," Computer, 45:4 (2012): 6-8

Afternoon Session II (2:30 p.m. to 3:45 p.m.)

Instructor: Dale Poulter

- Introduction to Amazon Web Services

Open session (4:00 p.m. to 4:30 p.m.)
- Participant presentation

Warren Center Reception (4:30 p.m. to 5:30 p.m.)
- The Robert Penn Warren Center for the Humanities will host a reception for the Institute at the Vaughn Home on the Vanderbilt Campus.

Dinner (5:45 p.m. at The Commons)

Thursday, June 19

Breakfast (8:00 a.m. in Rand Hall)

Morning Session I (9:00 a.m. to 10:15 a.m.)

Instructor: Kevin Clarke

- Architectural patterns for building larger applications
- Introduction to the Model-View-Controller pattern

Morning Session II (10:30 a.m. to 11:45 a.m.)

Instructor: Winona Salesky

- Introduction to eXist's XQuery Update Extension

Lunch (11:45 a.m. in Rand Hall)

Afternoon Session I (1:00 p.m. to 2:15 p.m.)

Institute Faculty

- Building a simple web search engine for the Shakespeare corpus – Part One

Afternoon Session II (2:30 p.m. to 3:45 p.m.)

Institute Faculty

- Building a simple web search engine for the Shakespeare corpus – Part Two

Open session (4:00 p.m. to 5:00 p.m.)

• Participant presentation

Dinner (5:45 p.m. at The Commons)

Friday, June 20

Breakfast (8:00 a.m. in Rand Hall)

Morning Session I (9:00 a.m. to 10:15 a.m.)

Instructor: Dale Poulter

• Deploying our Shakespeare search engine to the "cloud" using Amazon Web Services – Part One

Morning Session II (10:30 a.m. to 11:45 a.m.)

Instructor: Dale Poulter

• Deploying our Shakespeare search engine to the "cloud" using Amazon Web Services – Part Two

Lunch (11:45 a.m. in Rand Hall)

Afternoon Session I (1:00 p.m. to 2:15 p.m.)

Instructor: Kevin Clarke

• Review of XQuery fundamentals and questions

Afternoon Session II (2:30 p.m. to 3:45 p.m.)

Instructor: Kevin Clarke

• Review of eXist fundamentals and questions

Open session (4:00 p.m. to 5:00 p.m.)

Instructor: Clifford B. Anderson

• Institute wrap up
• Survey
• Recommendations for developing a stronger community of XQuery experts in the digital humanities

Dinner (5:45 p.m. in Rand Hall)

Appendix Two

VANDERBILT ✺ UNIVERSITY

research
news@Vanderbilt

ABOUT   RESOURCES   SUBSCRIBE   VU NEWS

Vanderbilt to host NEH-supported institute on digital humanities

by Ann Marie Deer Owens | Posted on Monday, Jun. 2, 2014 — 11:41 AM

Want more research news from Vanderbilt? Subscribe to our weekly newsletter »

Scholars from across the United States and abroad will be at Vanderbilt University June 9-20 to improve their skills in building projects in the digital humanities.

The XQuery Summer Institute: Advancing XML-Based Scholarship from Representation to Discovery will be led by Clifford B. Anderson, director for scholarly communications at the Jean and Alexander Heard Libraries at Vanderbilt.

During the two-week session, 12 participants will learn how to build projects in the digital humanities using the programming language XQuery and an open source XML database called eXist.

"Our goal is to help digital humanists take the next step in their work," Anderson said. "We want them to reach the full scholarly potential of encoding documents in machine readable format by teaching them how to query and analyze those texts computationally."

The institute, hosted by Vanderbilt, is among the Institutes for Advanced Topics in the Digital Humanities supported by the Office of Digital Humanities at the National Endowment for the Humanities.

Anderson (Vanderbilt University)

Jody Combs, associate dean for the Heard Libraries, said that hosting the institute underscores Vanderbilt's commitment to digital scholarship. "As new modes of computation-driven inquiry emerge in the humanities, we need to cultivate and to provide support for new kinds of humanities scholars," Combs said. "This award from the Office of Digital Humanities helps build capacity for digital research in the humanities locally at Vanderbilt as well as nationally and internationally."

David Michelson, assistant professor of the history of Christianity and classics at Vanderbilt, and Dale Poulter, coordinator of search and core services in the Heard Library, will serve as project faculty. They will conduct instruction in consultation with Jonathan Robie, one of the inventors of XQuery, and two other expert consultants.

"Simply put, XQuery is a programming language designed from the ground up to work with XML," Anderson said. "While XQuery excels at querying, manipulating and remixing XML documents, it can also be used like a general purpose scripting language." An interview with Anderson about the XQuery Summer Institute is available on the NEH website.

Vanderbilt University is also host to an NEH-funded digital humanities project for the study of minority culture in the Middle East and Asia – Syriaca.org: The Syriac Reference Portal, directed by David Michelson.

Contact:
Ann Marie Deer Owens, (615) 322-NEWS
annmarie.owens@vanderbilt.edu ✉

Share This Story

f ⚫ ⚫ ✉ 🖨 ➕ 8

Tweet 41   Like 8   +1 0

Explore Story Topics

📁 myVU, myVU News, releases, Research, Society and Culture ⬦ Arts and Science, Clifford Anderson, Dale Poulter, David Michelson, digital humanities, divinity, featured research, Jody Combs, library, National Endowment for the Humanities, NEH

Related Stories

Learn about podcasting, makerspaces, more at Library workshop series JANUARY 20, 2015

Two Vanderbilt researchers awarded NEH grants DECEMBER 15, 2014

Vanderbilt Central Library to host second Wikipedia 'edit-a-thon' OCTOBER 29, 2014

Subscribe

There are lots of ways to keep up with Vanderbilt research news. Choose your preferred method:

Sign up for the weekly Research News @Vanderbilt e-newsletter.

Name
Email

Subscribe to Research

More subscription options »

Latest Research News

A VU Inside: Engineer uses cotton candy to build artificial blood vessels MARCH 27, 2015

Engineering dean joins nationwide pledge to train next-gen engineers MARCH 27, 2015

'College of the future' topic of engineering

19